Discourse Analysis Techniques for Modeling Group Interaction⋆

Alexander Feinman and Richard Alterman

Computer Science Department, Brandeis University MS018,
415 South St., Waltham MA 02454
[afeinman,alterman]@cs.brandeis.edu

Abstract. This paper presents discourse analysis techniques for modeling the interaction of a small group of users engaged in same-place / different-time interaction. We analyzed data from use of *VesselWorld*, our experimental testbed, and formulated a modeling technique based on the recurrence of coordination problems and the structure that users create to handle these problems. Subsequent experiments revealed that this analysis had failed to capture issues with the cognitive load required to maintain common ground. By tagging the references users make to both domain and conversational objects, we were able to extract patterns of information access and model the cognitive load incurred to maintain common ground.

1 Introduction

Groupware applications for supporting different-place interaction can be very hard to use, even impairing the work they are built to support. Design of successful groupware applications is a challenging problem for many reasons, ranging from technical challenges to sociological issues of introducing groupware into existing interactions. Ideally, a software system is built to support the practice of work that emerges in a community of users. For groupware systems, this requires modeling the interaction of users as they perform their work.

This paper presents discourse analysis techniques that model group interaction, and presents results based on data collected using our groupware testbed, *VesselWorld* [7]. Initially, we performed an analysis that looked at communication to model recurrent issues of coordination and the conversational structure users created. VesselWorld data indicated that we needed to model the cognitive load of maintaining common ground. To examine how users were maintaining common ground, we created a second technique to model the referential structure users created. By building a model of how users published, accessed, and modified information, we determine how a system should mediate that information to improve performance. In particular, the length of time that a topic is relevant, the frequency of reference to the topic over that time, and the pattern of

access to the information all serve as useful indicators of cognitive load. Finally, the modeling method is used to explain observed results from the VesselWorld experimental data.

2 Modeling Recurrence in Coordination

To study issues in same-time / different-place coordination [4], our lab constructed the VesselWorld groupware system. Three users conduct a simulated clean up of a harbor containing toxic waste, which must be moved safely onto garbage barges by the users. As the users interact, the system logs all actions and communication for later analysis. This makes VesselWorld ideal for exploring issues of group interaction. At this time, we have collected over 250 hours of usage data using the system.

Each of the three users captains a ship navigating the harbor: two users pilot ships with waste-retrieval cranes attached (referred to as crane1 and crane2), and the other pilots a tugboat (referred to as tug1) able to move the some barges around the harbor. The tugboat is also able to identify waste, and seal the leaks caused by mishandling of waste. The harbor is cleared in a turn-based fashion, with each user explicitly planning actions before submitting them to the system for evaluation. During a session, the users are physically separated, but are able to communicate freely via a textual chat facility built into the VesselWorld system.

1. crane2: ok, then XLD right?
2. crane2: sub Lift
3. tug1: yep
4. crane1: k
5. crane2: sub load
6. crane1: k

Fig. 1. Dialogue from the VesselWorld system.

Despite the simplicity of the harbor-clearing task, users of the VesselWorld system had extensive problems staying coordinated. Problems with maintaining common ground, communicating intentions, and managing information appear throughout the data. To handle these issues, users created conversational structure. For example, to lift and load large barrels of waste, the two cranes must submit a `lift` command on the same turn. This led to the appearance of adjacency pairs, as seen in Figure 1, to ensure that users were in fact lifting on the same round.

Given such data, we turned to discourse analysis to help make sense of it. Starting with possibly the most salient form of human interaction, conversation, researchers have been able to formulate theories about rules governing conversational flow [8], grounding in communication [3] and the purpose of conversational

acts performed by users [9]. We developed an analytic method that built on this research that searches the discourse for three indicators that adaptation of the system might be advisable:

1. Recurrent patterns of coordination.
2. Repeating errors in coordination.
3. Development of conversational structure by the users to organize recurrent problems of coordination.

Recurrent patterns of coordination and repeating errors are situations in which the users might introduce conversational structure to better organize activity so as to improve performance. The last indicator goes a step further; in this case, the users have determined a potential area of improvement, and have also devised structure to simplify their conversation about the situation and thereby improve it. Examples of structure generated by users include the invention of notational conventions like specialized jargon ("XLD" meaning an extra-large waste, dredge equipment required), and conversational procedures that support the coordination of activities like adjacency pairs, as seen in Figure 1.

Of the three indicators, the third seems to be the surest bet for the analyst; the users have added corroborating evidence for a particular analysis of the situation. There are problems, however, with an analysis strategy that relies exclusively on the third indicator. The data would have to be extensive, so the users would have time to generate all the most useful secondary structures. The users might identify a coordination problem they would like to fix, but lack the means to fix the problem – there is some evidence for this in our data. Finally, there is no guarantee that organizational structure users might add would improve the situation at all.

2.1 Applying the model

Based on analysis of data generated in our VesselWorld pilot study, three *coordinating representations* (CRs) – software tools for simplifying context sharing – were introduced into VesselWorld: the *shared planning* window, the *object list*, and the *high-level planning* window. These representations were aimed at capturing specific forms of information being passed between the users: fine-grained timing issues in shared planning, coarse-grained planning and ordering issues in high-level planning, and object references in the object list.

One of these, the object list is pictured in Figure 2. It presents information about barrels of toxic waste in the harbor in a concise tabular format. A user adds an entry to the object list by entering details about it in the entry fields at the top and clicking "Add Entry". The entry then appears in the table at the center of the object list, which is visible to all users. Entries can be modified in place by clicking on the cells of the table; updates are sent to all users.

Fig. 2. A coordinating representation: the object list.

2.2 Results

To evaluate the coordinating representations we ran the *VW3* experiment, completed in early 2001. The experiment was a single-variable study conducted to verify the utility of these coordinating representations. Six groups of three people were paid to use VesselWorld for 12 hours each, including training time. Over the course of the experiment, the performance of user groups improved dramatically, reaching a plateau of sorts only after the fifth or sixth hour of use; as a result, we examined only data from the last five hours for each group, in an effort to avoid interference from the effects of having novice users. Three groups used the base (*non-CR*) version of Vesselworld, and could only communicate via textual chat. The other three groups used a version that included the three CRs. Differences were dramatic; users of the CR system completed similar problems in 49% of the clock time that users of the non-CR system required, and committed 61% fewer errors. A complete summary of experimental results can be found in [1].

Significantly, users did not make full use of the provided CRs. Portions of the object list dealing with waste status went unused, or were used very rarely; and one of the three CRs, the high-level planning window, was not used at all. To figure out what had gone wrong, we developed an analytic method based on discourse analysis that examined the referential structure of the users' discourse.

3   Modeling Cognitive Load

We noted that many of the issues we saw with the existing CRs could be attributed to imbalance in the collaborative effort required to use those CRs. To examine these issues, we created an analysis method centered around the sharing of information between users. In a shared environment like VesselWorld, users interact by negotiating over both domain objects and conversational objects. Insight from the initial analysis indicated that there was utility in quantifying and tracking these exchanges of information.

Preliminary analysis of information exchange in the VesselWorld data showed that information of different categories was handled quite differently. Waste and

other information relating to objects in the domain tended to have a long lifetime of relevance (time between first mention and last mention), and the users accessed this information infrequently but steadily during that time. Plans, on the other hand, dominated conversation while they were being discussed, but rarely were reused or referred to once agreed upon unless an error occurred. Repairs, where users attempted to fix mismatches in common ground, correct errors in the interaction, or disambiguate misunderstandings, dominated conversation until corrected. Only occasionally would another message find its way into the thread of conversation while a repair was underway.

In VesselWorld, information usually has a simple life cycle. Information is reported by a user, either in the chat window, or (in the CR groups) via the object list or shared planning window; it may then be noted by other users; it again becomes relevant at some point and must be retrieved. Information that is relevant over a long span of time may be updated, modified, or otherwise accessed before it is finally rendered irrelevant and forgotten. While information may reside in the short-term memory of the users, storing complex information there can put an unreasonable cognitive burden on the user; distributing the information into the environment [6] yields superior results. However, distributing too much information into the environment raises issues of user confusion, and may create unnecessary work on the part of the designer.

It is important, therefore, to determine what information is worth supporting by constructing CRs. We formed two main hypotheses:

1. Information with a *long period of relevance* is worth recording in a external representation
2. Information which is *accessed frequently* by the users is worth recording in an external representation

These hypotheses serve as general indicators of when adapting the system may simplify coordination. In the first case, information that is relevant over a long period is very likely to be irrelevant for some subsection of that period. During this time, it represents an unnecessary burden on the user's short-term memory, and in a complex situation with many such items, the burden can easily outstrip the user's ability to memorize. In the second situation, the information is used frequently, and hence needs to be readily available to the users. However, relying on having this information occupy the short-term memory of the users is as unrealistic in this situation as in the previous one. Hence, it needs to be easily accessible to users as they perform their tasks.

3.1 Identifying referential structure using iotas

In order to highlight information falling into the above categories, an analyst must track referents users refer to by tagging the discourse. The tagging scheme consists of identifying the information that users share and tracking its subsequent use within the system. Because we are interested in tracking both conver-

sation and task structure, we must examine referents of both types. To avoid the tongue-twister "references to referents", we will call these referents "iotas".

An iota represents a simple conversational item that the users refer to. Examples include: a barrel of toxic waste in the VesselWorld domain; a plan to clear a barrel of waste; a repair of differences between common ground between two users; a conventional procedure for handling a particular situation. In general, any sort of information that the users refer to qualifies. To make this clear, let us follow through an example taken from a non-CR group in their final VesselWorld session.

```
      ⋮
 7. tug1: mX at 400 125                    [IOTA-7 waste: mx@400,125]
 8. crane1: medium at 392 127              [IOTA-8 waste: m?@392,127]
 9. crane1: that's got to be the same one  [IOTA-9 repair: IOTA-8 same as IOTA-7]
10. tug1: yep                              IOTA-9
11. tug1: that's an mX                     [IOTA-7 waste: mx@392,127]
```

Fig. 3. Analysis of discourse from a non-CR VesselWorld session

The analysis shown in Figure 3 looks at a short segment from a non-CR group. The users have encountered a few wastes, and are sharing what they see so that they can plan how to clean up the harbor. First, in line 7, the user operating the tug reports information about a nearby waste, using an established shorthand: "mX at 400 125". Here, mX indicates that the tug is talking about a waste of medium size (m) that requires no special equipment (X) to handle safely. The tug indicates that the waste can be found at the location (400, 125) in the harbor. To create a waste iota for this referent, the analyst generates a unique IOTA name based on the current line of discourse (IOTA-7), puts the IOTA tag in square brackets to indicate that this is a new or modified reference, denotes the type of the iota as "waste", and lists all information available about that waste.

In line 8, crane1 almost simultaneously reports on a waste nearby. The analyst again creates an iota to track it. Here, the type of equipment needed is unknown to the user, as only the tug can ascertain equipment needs. The user indicates this by omission; the analyst instead uses a question mark in the iota definition.

In line 9, crane1 notes the similarities between the two waste reports: both wastes are medium-sized, and they are located very close to each other. Likewise, the reports of equipment (unknown vs. none) are not contradictory. Perhaps most telling, crane1 can likely see the area the tug is referring to, and does not see a waste there. The users have run into this situation before, and crane1 quickly proposes (in line 9) a need for repair of common ground. The analyst notes the repair as an iota of type "repair", and makes sure to mention the iotas that are involved.

The tug, who can also see both locations (400,125 and 392,127), and is able to refer to the info window to get the exact coordinates for the waste, agrees to

the repair in line 10. It appears that the tug estimated the original specification of the waste location (400,125), rounding to the grid intervals visible on the user's display. The analyst notes the agreement to the repair, and refers to the already-instantiated iota IOTA-9 by naming it without square brackets.

The tug reviews relevant information about the waste in line 11. This acts as evidence supporting the repair (that the two references refer to the same waste). The analyst updates IOTA-7's expansion (again using square brackets, this time to indicate that the contents of the iota are being modified), and chooses the earlier of the two names for the waste (IOTA-7 and IOTA-8) to disambiguate further references to the waste.

3.2 Results

A great deal of information can be gleaned about group interaction and the passing of information from tagging the discourse. However, in a long session, it is hard to keep track of the large number of iotas that are generated. To help with the task of interpreting the data we developed a number of special-purpose visualization tools. The access pattern for each iota – that is, when and how it is referred to – can be visualized using a program that extracts the iotas from the analysis, summarizes information about the iotas, and graphically represents the access pattern for each iota as a timeline.

Iota type	Mentions	Lifetime	Density
Vessel	3.1	183.6	1.7%
Waste	6.6	168.7	3.9%
Location	2.6	62.6	4.2%
Convention	5.5	109.5	5.0%
Barge	11.9	294.0	5.6%
Plan	3.4	12.0	28.5%
Repair	3.0	4.8	62.5%

Fig. 4. Selected results from analysis of the VW3 experimental data.

Some results from analyzing the VW3 data are summarized Figure 4. *Mentions* counts the average number of times an iota is mentioned in any user utterance; for example, in Figure 3, IOTA-8 is mentioned twice. *Lifetime* is the average number, inclusive, of utterances between first and last mention; for IOTA-8, 2. Density is the ratio of these two numbers, and represents the percent of utterances over the life of an iota that refer to that iota; as IOTA-8 dominates conversation over its short life, its density is 100%. These measures are calculated for each iota, and the average for each type is calculated.

Raw data about lifetime of relevance, frequency of reference, and density of reference can be represented in, for example, a scatter plot, to highlight clusters of iotas with similar information access patterns. In Figure 5, the domination
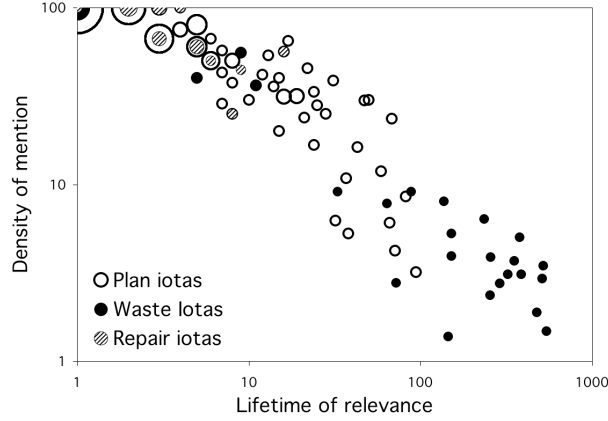
Fig. 5. Distribution of iotas for a non-CR group (larger bubbles indicate multiple iotas)

of longer lifetimes by waste iotas is made quite visible. Iotas in the lower-right quadrant of the scatter plot represent those that will require the most long-term coordination work to maintain common ground. Those in the upper-left quadrant represent those that by their nature will dominate conversation. Using these and other techniques, we drew conclusions about the workload required to stay coordinated over each sort of information. Here we present results for the most frequent types of iotas used by users in the VW3 experiment, plans and wastes, which together comprised over 80% of the iotas seen.

Plans As can be seen in Figure 5, plan iotas generally have a short lifetime of relevance, measured as the total number of utterances from the first mention of the iota to the last mention: 85% of plans seen had a lifetime of 6 or fewer utterances. During this short time they had a high frequency of occurrence, with 25% or higher density in 90% of the plan iotas seen, and 50% or higher density in 70% of the plans seen. Only in about 10% of the cases will a plan drag on for some time, generally when it is postponed in favor of another plan; it then shows a pattern resembling two plans stuck together, one starting at the initial negotiation, and one starting at the (postponed) re-negotiation.

This distribution of most plans leads us to conclude that, while negotiation of a plan is crucial, their short lifetime means that users will be unlikely to want to put any effort into constructing a shared external representation of their plan. In addition, because a small number of plans co-exist at any one time, and they tend to dominate conversation while they are being discussed, users are probably storing them in short-term memory and using the recent chat for reference. Hence, while an external representation of a plan may serve to ground the plan for multiple users, it is unlikely that they would require persistent storage of plans.

This prediction agrees with the experimental evidence produced by the VW3 experiment. Our users refused to use the high-level planning window, which required significant user effort to create and share a representation of plans, but allowed persistent storage of these plans. However, they fully embraced shared planning, which allowed a low-impact (if less general-purpose) method for sharing short-term future actions. Because it provided good visibility of current intentions, coupled with a low cost to communicate those intentions, shared planning matched the short-term, non-persistent usage pattern of plans.

Shared domain objects The most common type of shared domain object, wastes, have a quite long lifetime of relevance: anywhere from 50 to 500 utterances, depending on session length, complexity, and the order in which the users clear the wastes. In addition, there are short periods of intense activity relating to the wastes, separated by long stretches (10-40% of the lifetime of an iota, in various cases) where they are not discussed at all. This periodic distribution imposes a heavy load on short term memory for the period of irrelevance (the long stretches where the waste information is not accessed). Also, because of the long lifetime of relevance for wastes, more than one waste is usually relevant at any one point in time; hence, when retrieving information about wastes, users will first need to choose a waste. Hence, storing waste information requires a persistent representation providing easy access and modification capabilities.

In the non-CR system, we provided two methods for discussing wastes: chat history and private markers. As waste is discovered, users generally announce their discovery in chat. Other users could access waste information in the history of their chat, but this was difficult, it would be interleaved with all other communication, and so most users used private markers instead. These allowed a user to annotate his or her map view with a simple note, akin to a sticky note. They were used extensively by most users to keep track of the waste. But markers had a number of problems. First, they effectively tripled the amount of group work needed to stay coordinated, by making each user place a marker in their own private environment. Second, placing a marker required the user to take information from chat, process it correctly, and place a marker in the appropriate position on the map. These complications were a source of great error in non-CR groups. Users often discovered mismatches in their sets of markers, prompting one group to introduce the "marker check", a complex conversational structure, to synchronize these private representations.

The object list (Figure 2), which consolidates and shares information about wastes and other domain objects, addresses these issues. The object list gives a shared, persistent representation of waste information in a form that users can easily access and parse. To support the periodic distribution characteristic of waste iotas it is important to construct a persistent and easy-to-access shared representation. Despite the relatively high cost of formulating an object in the object list (generally 15-30 seconds of interface work, mixed between mouse and keyboard), users willingly performed the task, perhaps sensing a return on the time investment: retrieval times for waste information ranged from five seconds

to over a minute when the user was forced to scan through the chat history, in comparison to the few seconds required for a quick visual scan of the object list.

4   Conclusions and Future Work

We have presented two discourse analysis techniques for modeling group behavior. First, we presented a method that looked at recurring problems in coordination, and the structure users constructed to handle them. However, this failed to take into account the cognitive load associated with maintaining common ground. We then introduced a second technique that tags user discourse to track the exchange of information by users, and discussed visualization tools for interpreting the data generated. From this analysis we are able to draw a picture of the workload required for users to stay grounded.

Building on work by Gray and Fu [5], Clark and Brennan [3], and others, it may be useful to look at quantifying the costs associated with communication for maintaining common ground. Users take a path of least collaborative effort when deciding what medium to communicate information. We saw evidence for this in (for example) the object list. Despite a high cost of formulating utterances in the object list, groups universally used it, due to the reduction in group effort to maintain common ground thereafter. Currently, we are investigating use of *reading-writing plans*, an expansion of *reading plans* [2], as a general method for quantifying the costs associated with communication in groupware systems.

References

1. Alterman, R., Feinman, A., Introne, J., Landsman, S.: Coordinating Representations in Computer-Mediated Joint Activities. Proceedings of 23rd Annual Conference of the Cognitive Science Society (2001)
2. Carpenter, T., Alterman, R.: A Taxonomy for Planned Reading. Sixteenth Annual Conference of the Cognitive Science Society (1994) 142-147
3. Clark, H.H., Brennan, S.E.: Grounding in communication. In Resnick, L.B., Levine, J., Teasley, S.D., editors, *Perspectives on Socially Shared Cognition* APA Press (1991)
4. Ellis, C.A., Gibbs, S.J., Rein, G.L.: Groupware: some issues and experiences. Communications of the ACM **34** (1991) 38-58
5. Gray, W., Fu, W.: Ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head: implications of rational analysis for interface design. CHI Letters **3** (2001) 112-119
6. Hutchins, E.: Cognition in the Wild. Cambridge: MIT Press (1995)
7. Landsman, S., Alterman, R., Feinman, A., Introne, J.: VesselWorld and ADAPTIVE. Technical Report TR-01-213, Department of Computer Science, Brandeis University (2001) Also presented as a demonstration at *CSCW-2000*
8. Sacks, H., Schegloff,, E., Jefferson, G.: A simplest systematics for the organisation of turn-taking for conversation. Language **50** (1974) 696-735
9. Searle, J.: Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press (1969)